

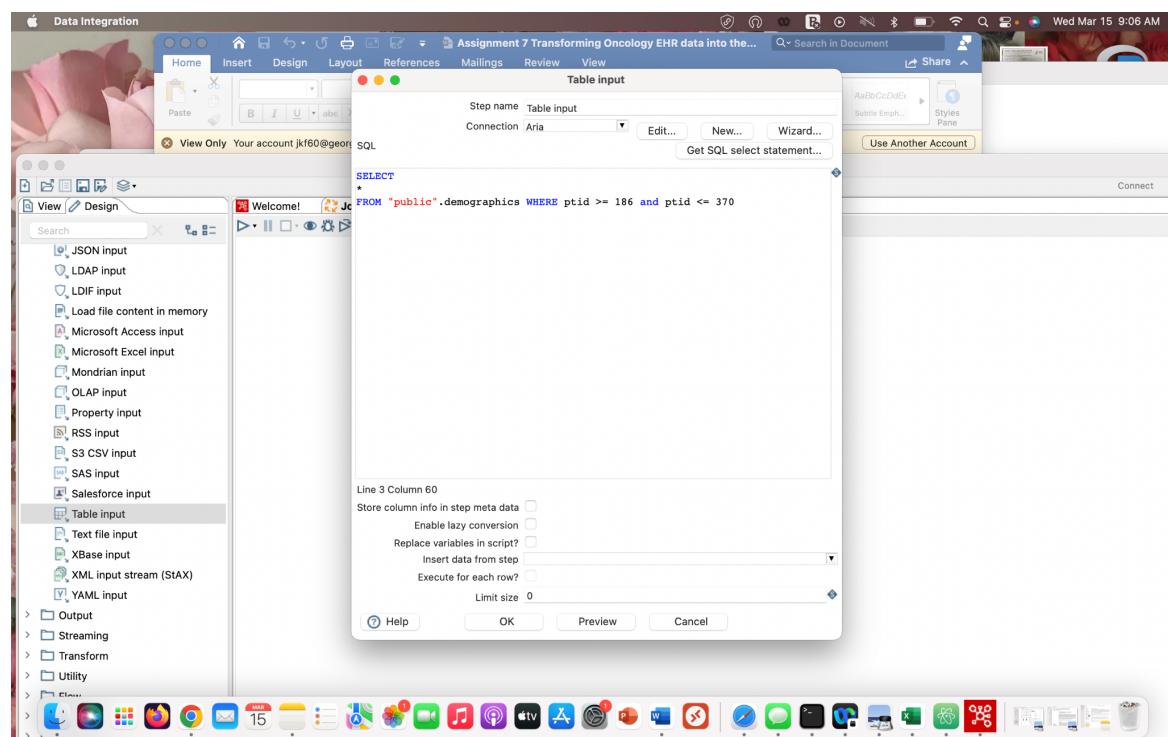
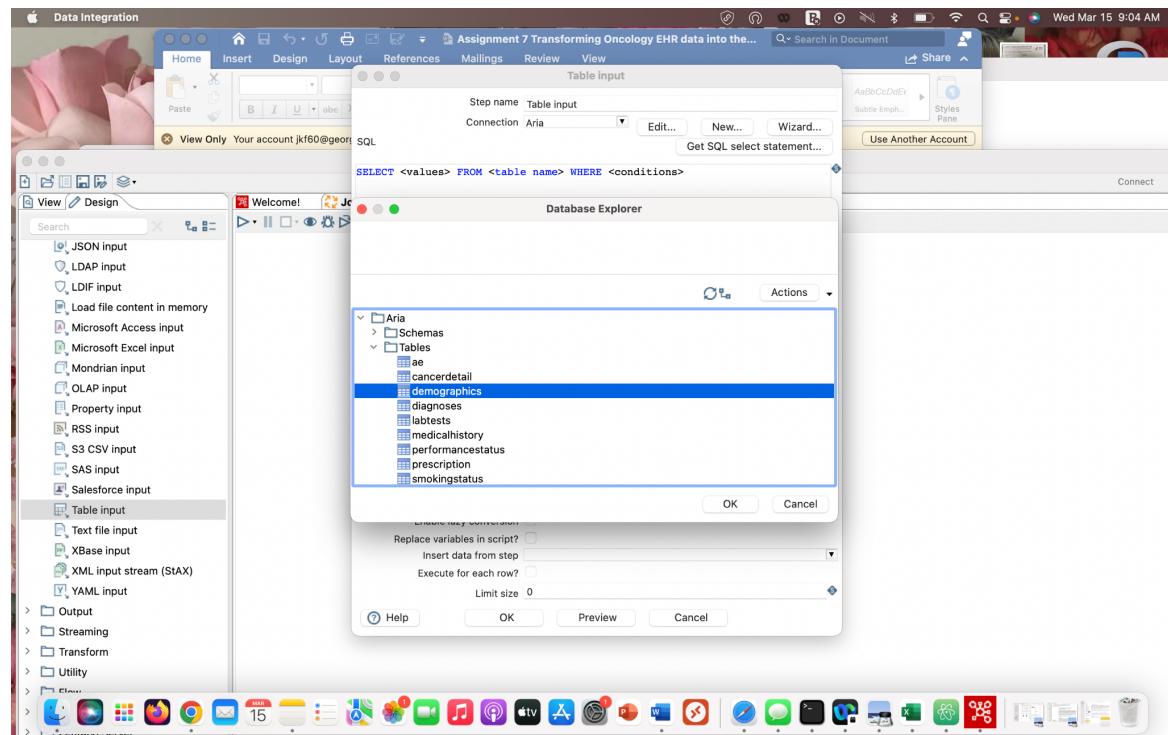
Joelle Fitzgerald

HIDS 505

Assignment 7: Transforming Oncology EHR data into the OMOP common data model

Part 1: Demographics ETL

Table Input



Preview of Table Input (selected patient range: 186-370 inclusive)

Rows of step: Table input (185 rows)

#	ptid	sex_cd	dob	educational_level	language	race_desc	ethnic_group_desc
1	186.0	M	1937.0	NULL	English	Black or African American	Not Hispanic or Latino
2	187.0	F	1937.0	NULL	English	Black or African American	Not Hispanic or Latino
3	188.0	F	1950.0	NULL	English	White	Not Hispanic or Latino
4	189.0	F	1962.0	NULL	English	Black or African American	Not Hispanic or Latino
5	190.0	F	1945.0	NULL	English	White	Not Hispanic or Latino
6	191.0	F	1939.0	NULL	English	White	Not Hispanic or Latino
7	192.0	M	1962.0	NULL	English	White	Not Hispanic or Latino
8	193.0	M	1958.0	NULL	English	Black or African American	Not Hispanic or Latino
9	194.0	F	1953.0	NULL	English	White	Not Hispanic or Latino
10	195.0	F	1959.0	NULL	English	White	Not Hispanic or Latino
11	196.0	F	1940.0	NULL	English	Black or African American	Not Hispanic or Latino
12	197.0	F	1968.0	NULL	English	Black or African American	Not Hispanic or Latino
13	198.0	M	1955.0	NULL	English	White	Not Hispanic or Latino
14	199.0	F	1983.0	NULL	English	White	Not Hispanic or Latino
15	200.0	F	1966.0	NULL	English	White	Not Hispanic or Latino
16	201.0	M	1964.0	NULL	English	White	Not Hispanic or Latino
17	202.0	F	1959.0	NULL	English	White	Not Hispanic or Latino
18	203.0	M	1945.0	NULL	English	White	Not Hispanic or Latino
19	204.0	F	1962.0	NULL	English	Black or African American	Not Hispanic or Latino
20	205.0	M	1940.0	NULL	English	White	Not Hispanic or Latino
21	206.0	M	1947.0	NULL	English	White	Not Hispanic or Latino
22	207.0	M	1957.0	NULL	English	NULL	Unknown
23	208.0	F	1962.0	NULL	English	White	Not Hispanic or Latino
24	209.0	F	1978.0	NULL	English	Asian	Not Hispanic or Latino
25	210.0	M	1952.0	NULL	English	Black or African American	Not Hispanic or Latino
26	211.0	F	1964.0	NULL	English	Black or African American	Not Hispanic or Latino
27	212.0	F	1953.0	NULL	English	White	Not Hispanic or Latino
28	213.0	M	1960.0	NULL	English	Black or African American	Not Hispanic or Latino
29	214.0	M	1969.0	NULL	English	Black or African American	Not Hispanic or Latino
30	215.0	M	1948.0	NULL	English	White	Not Hispanic or Latino
31	216.0	F	1971.0	NULL	English	White	Not Hispanic or Latino
32	217.0	F	1961.0	NULL	English	White	Not Hispanic or Latino
33	218.0	F	1949.0	NULL	English	Black or African American	Not Hispanic or Latino
34	219.0	F	1955.0	NULL	English	White	Not Hispanic or Latino
35	220.0	M	1964.0	NULL	English	White	Not Hispanic or Latino

Close Show Log

Rename to map Aria → OMOP

Spoon - Transformation 1 (changed) Connect

Step name: Rename

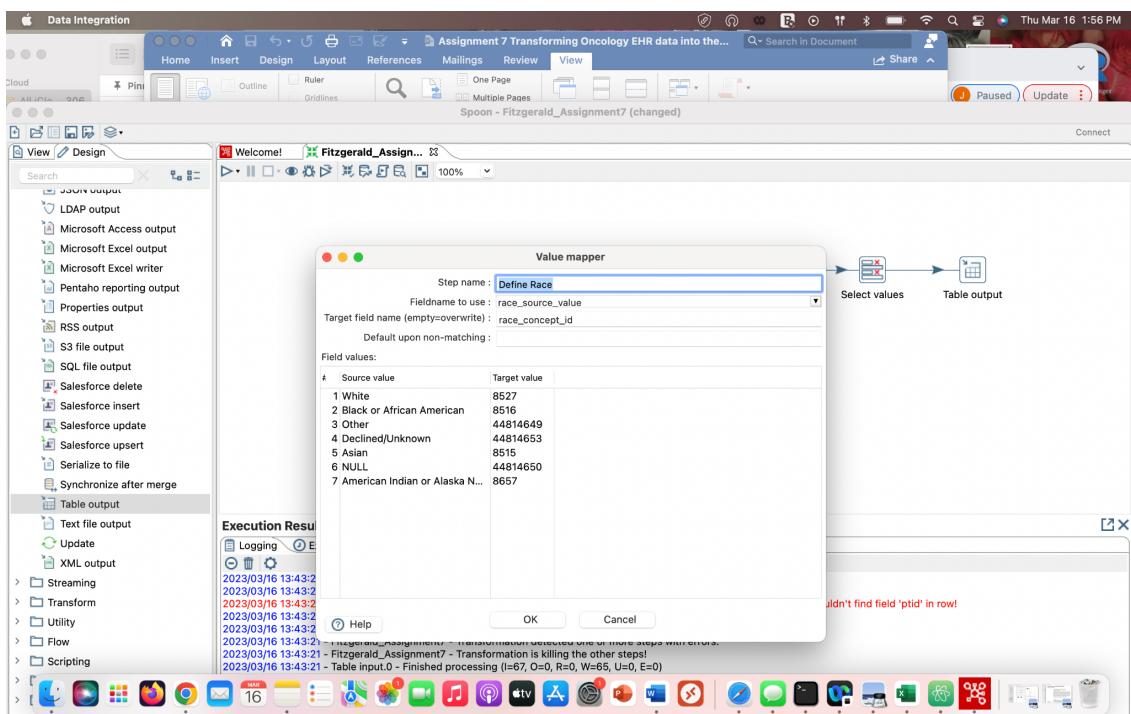
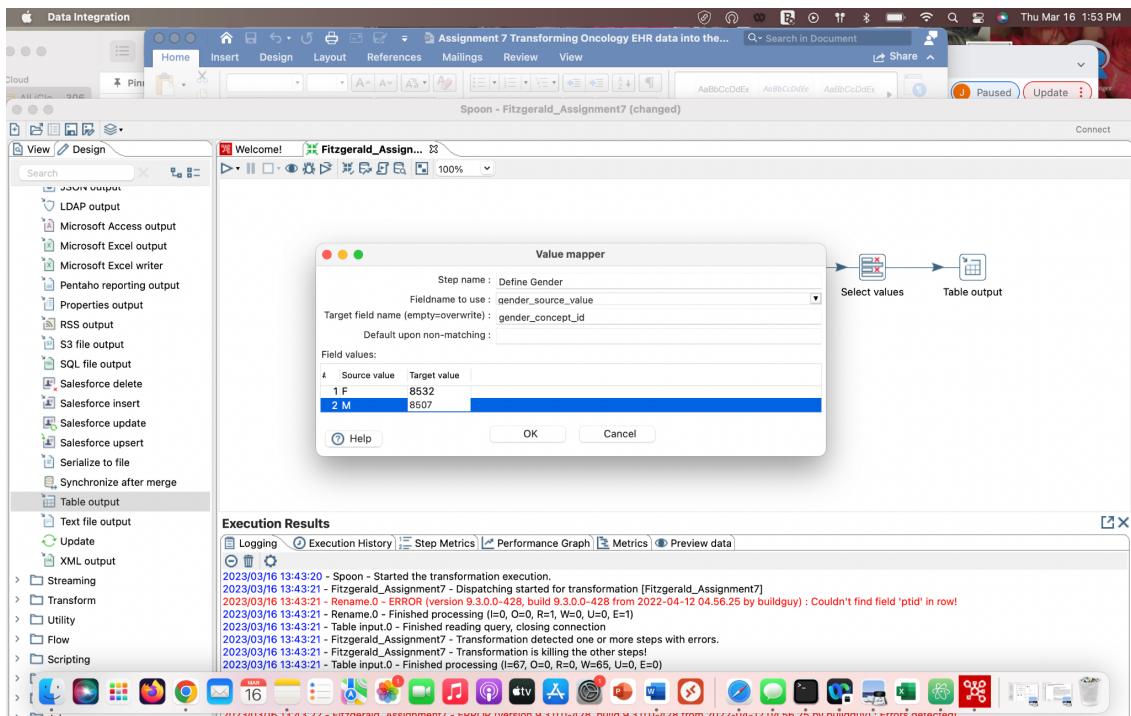
Select & Alter Remove Meta-data

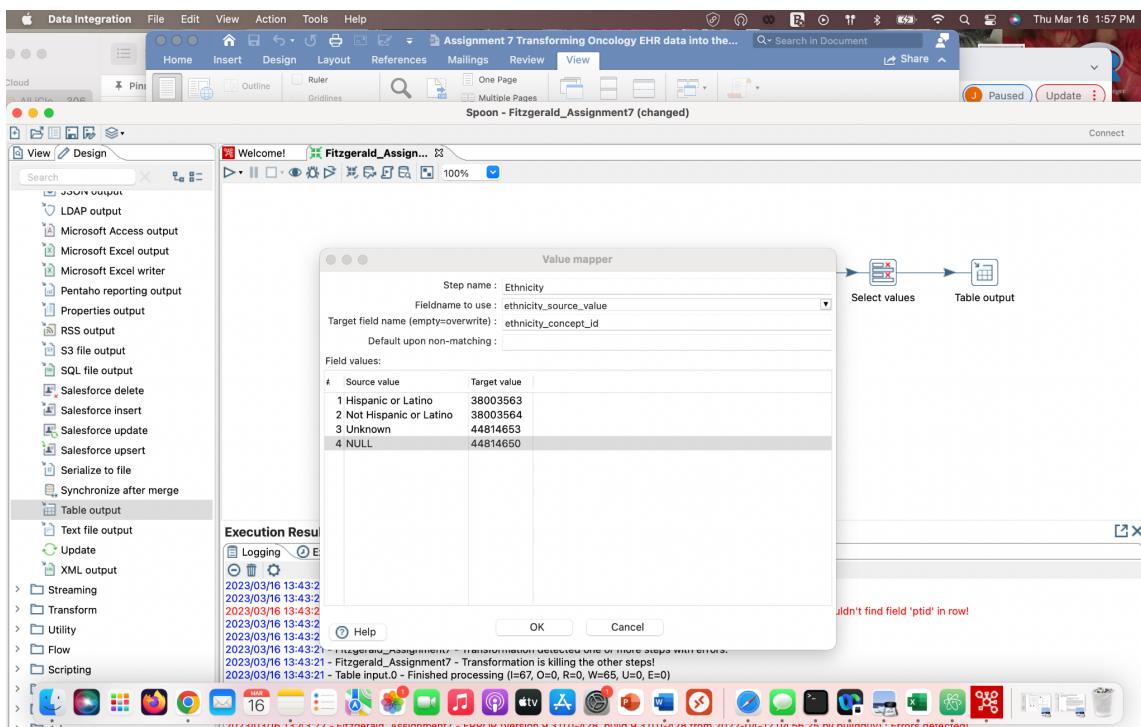
Fields :	Fieldname	Rename to	Length	Precision	Get fields to select	Edit Mapping
1 ptid	person_id					
2 sex_cd	gender_source_value					
3 dob	year_of_birth					
4 educational_level						
5 language						
6 race_desc	race_source_value					
7 ethnic_group_d...	ethnicity_source_value					

Include unspecified fields, ordered by name

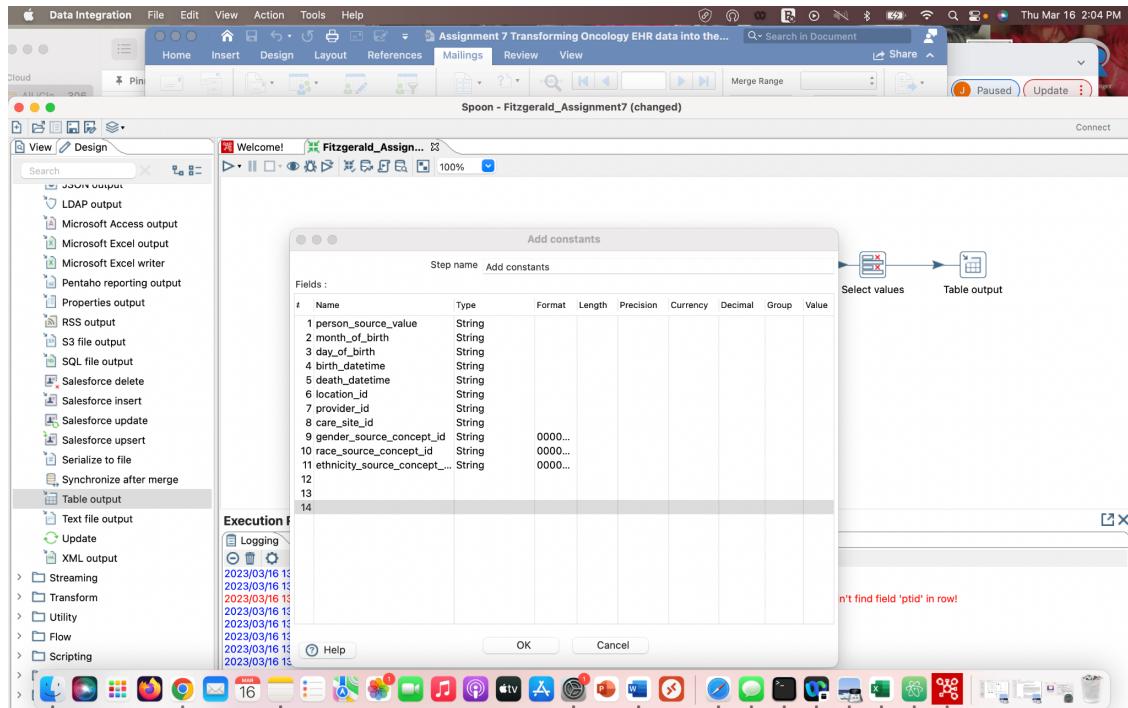
OK Cancel Help

Define Gender, Race, Ethnicity





Adding Constants



Selected Values

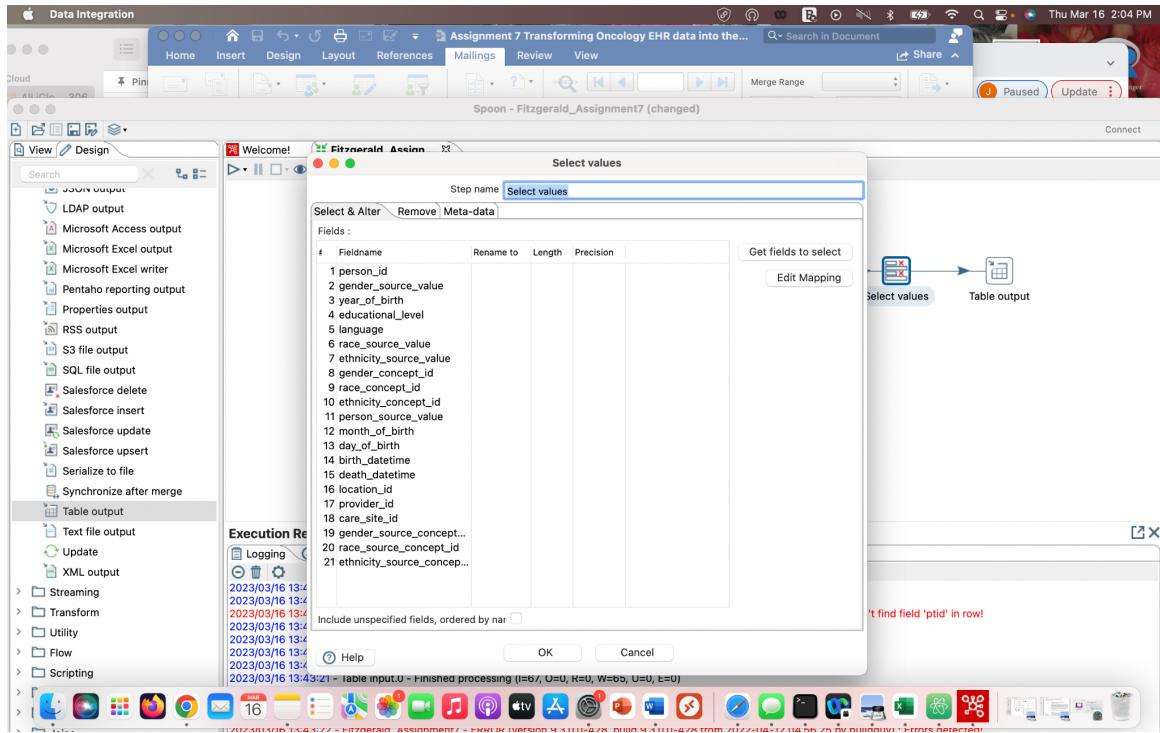
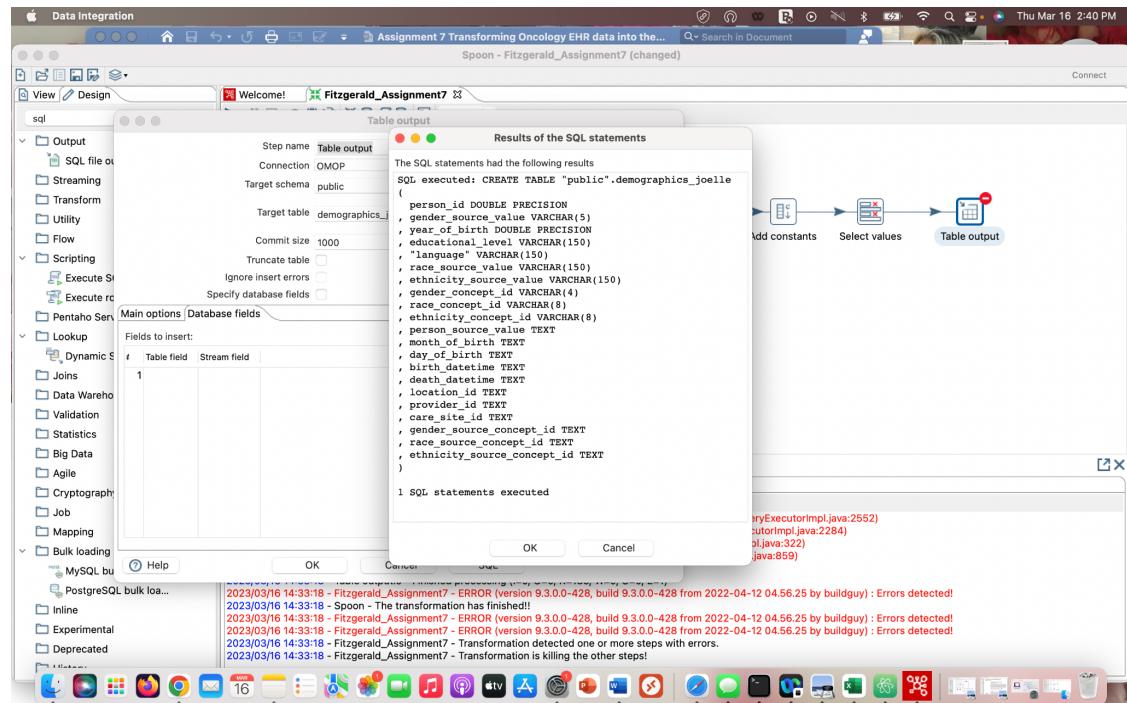
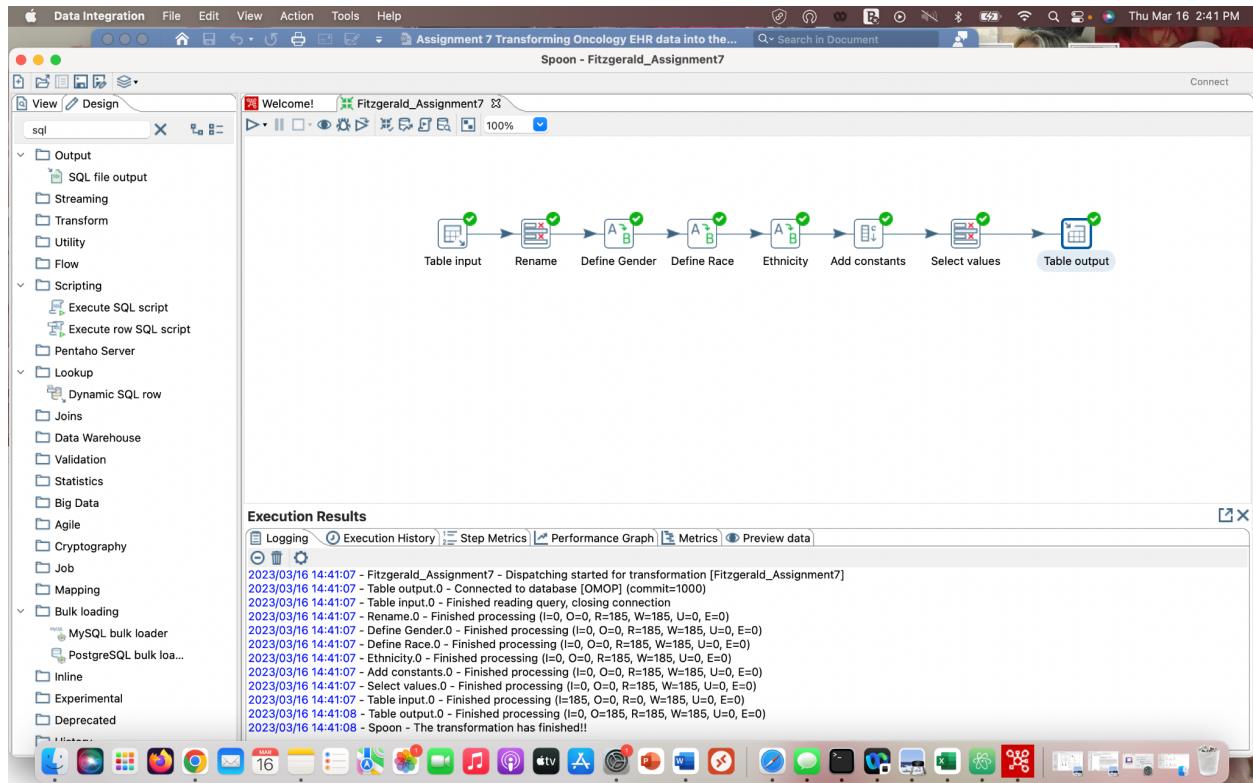


Table Output (OMOP)



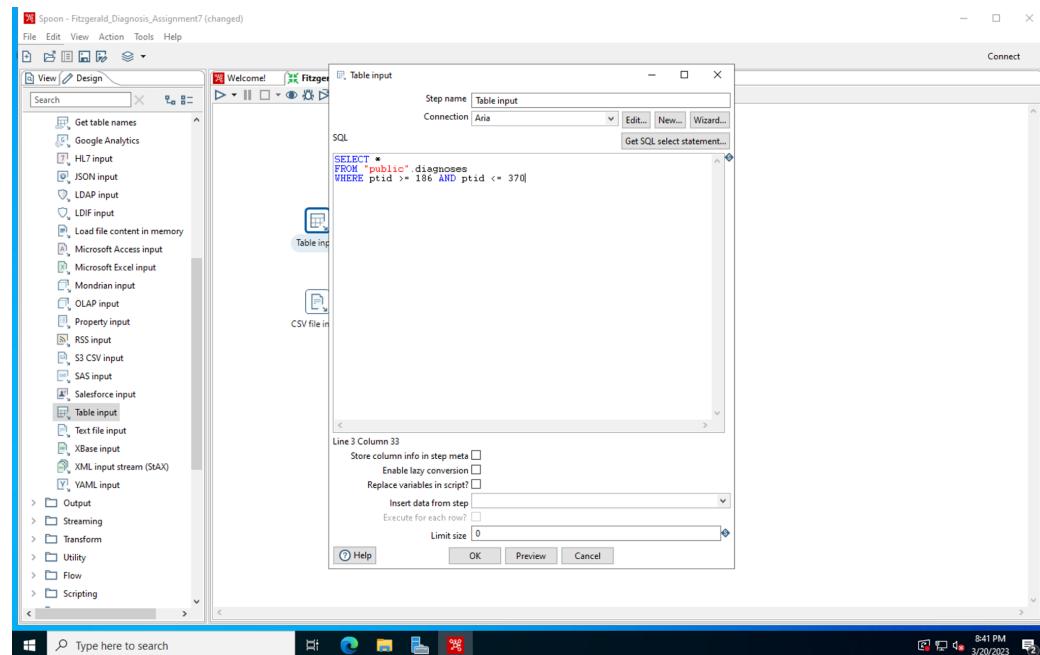
Successful Mapping



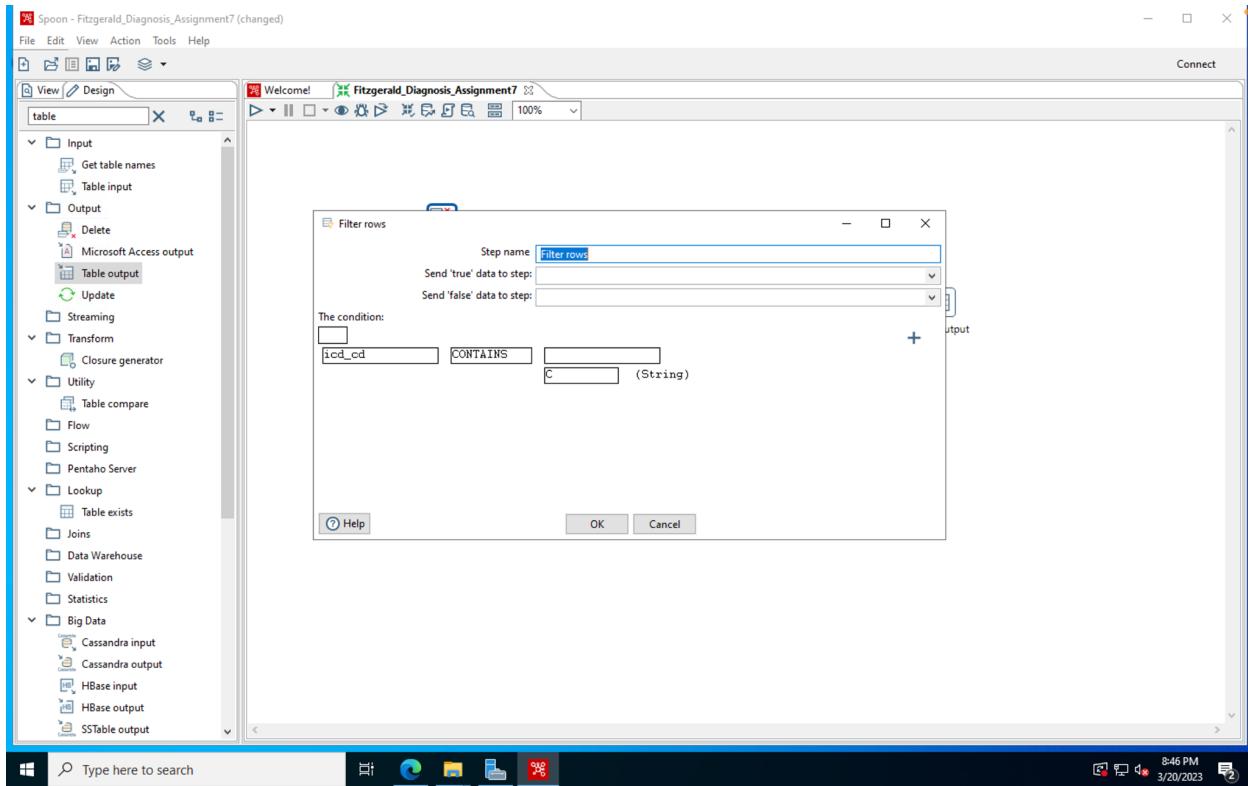
Part 2: Diagnosis ETL

(Some of assignment screenshots were done in VM)

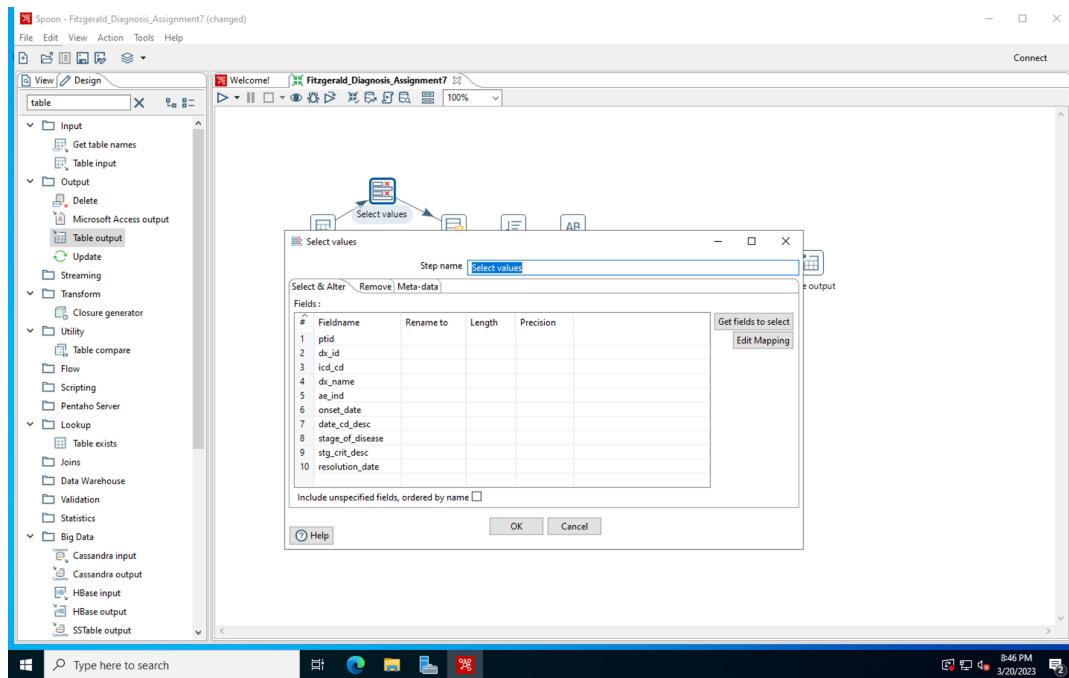
Table Input

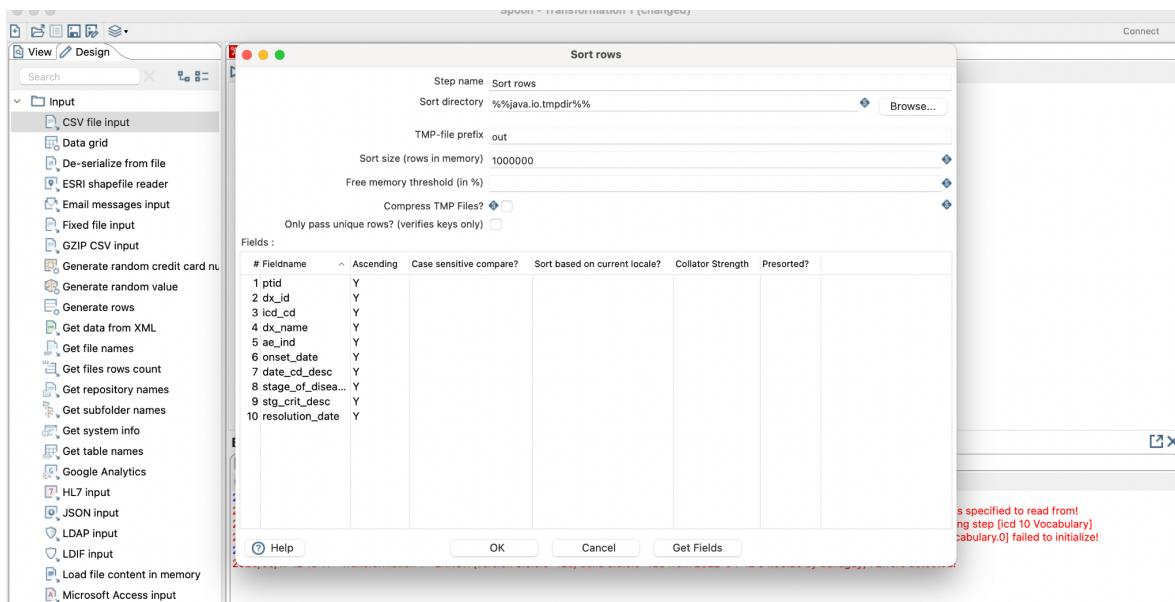
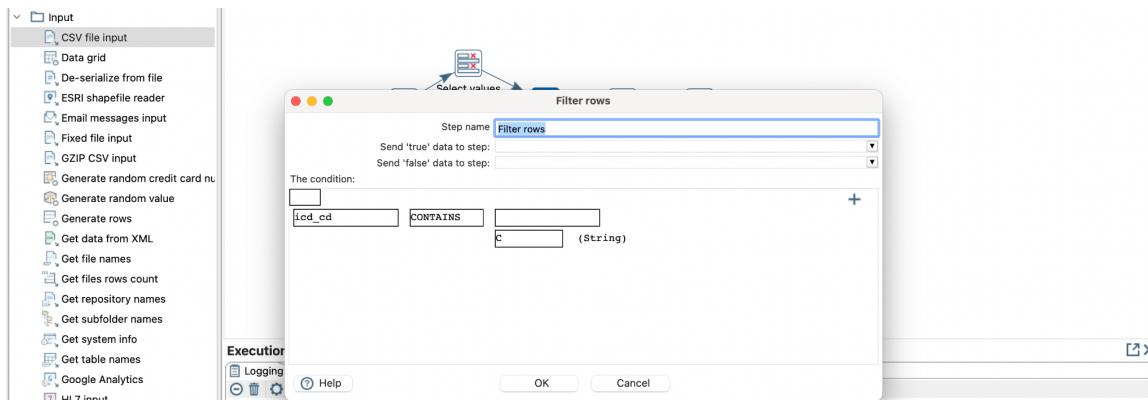
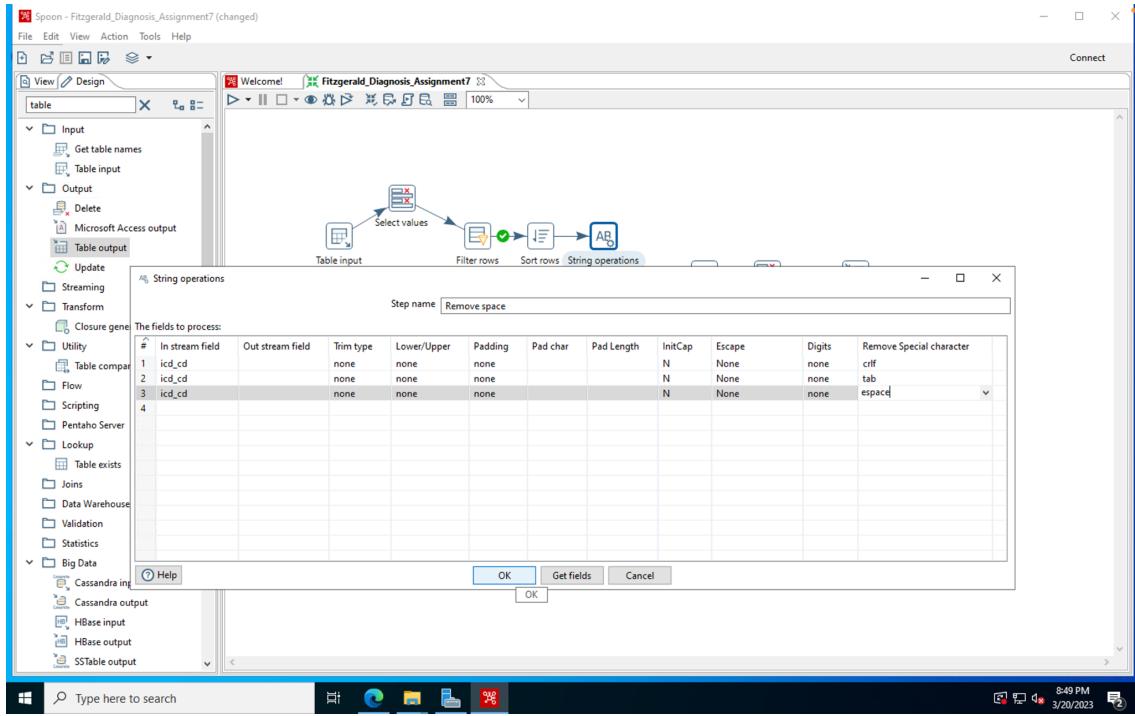


Filter Rows (based on cancer ICD10 diagnosis code id)

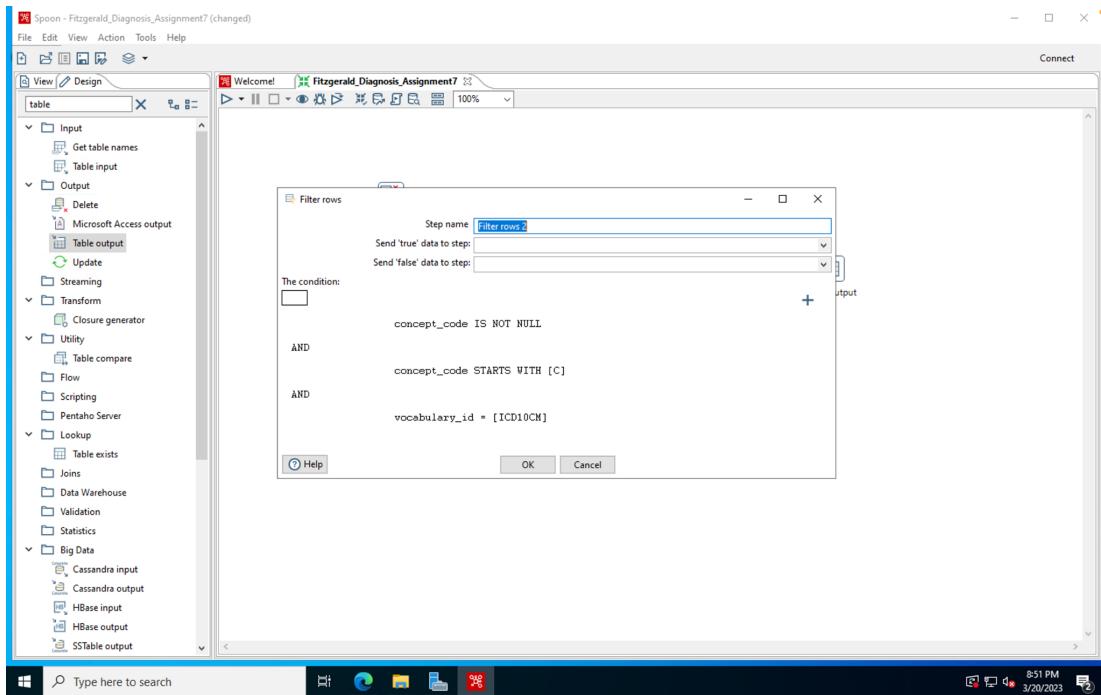


Selected Values → Filter → Sort Rows → Remove Space

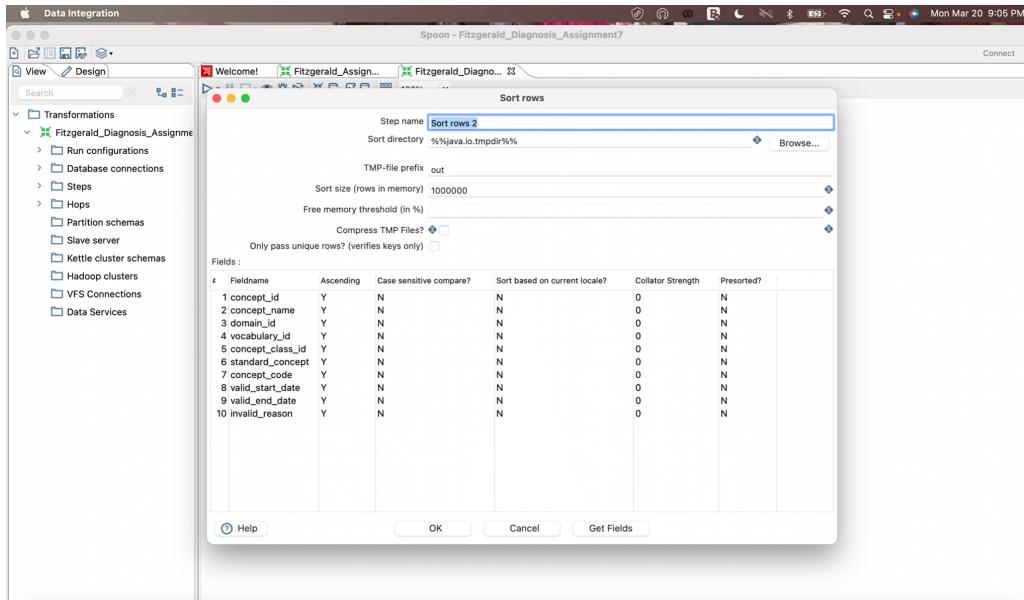


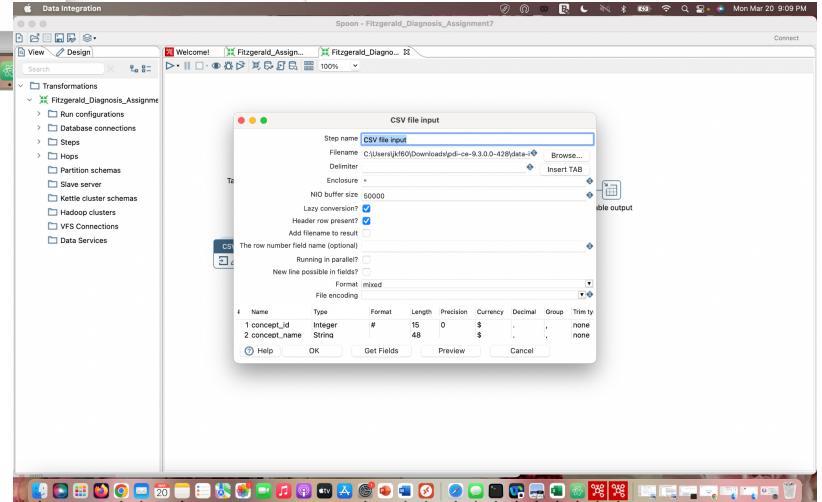
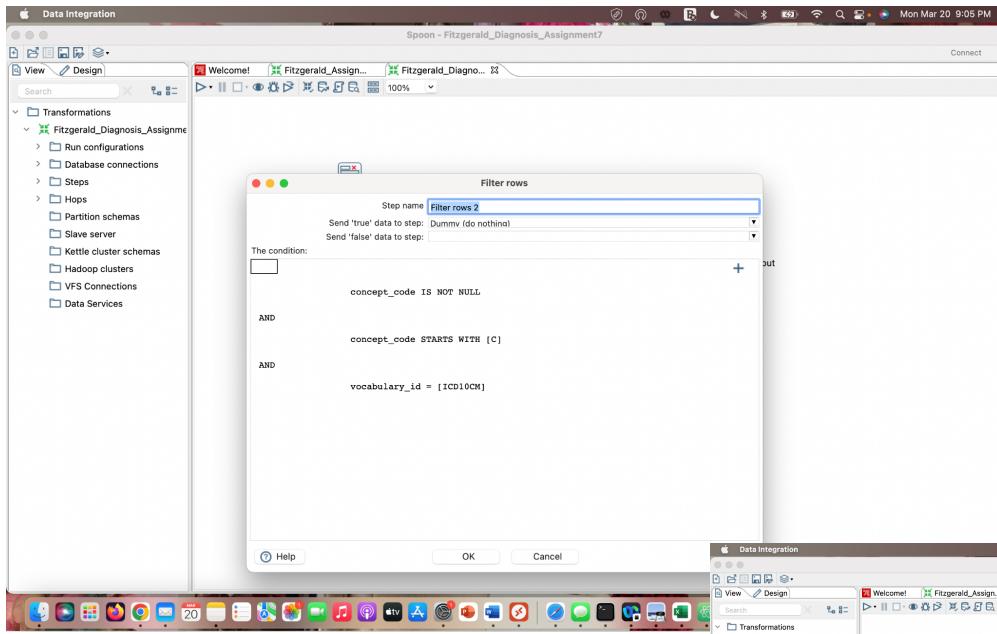


Filter Rows 2

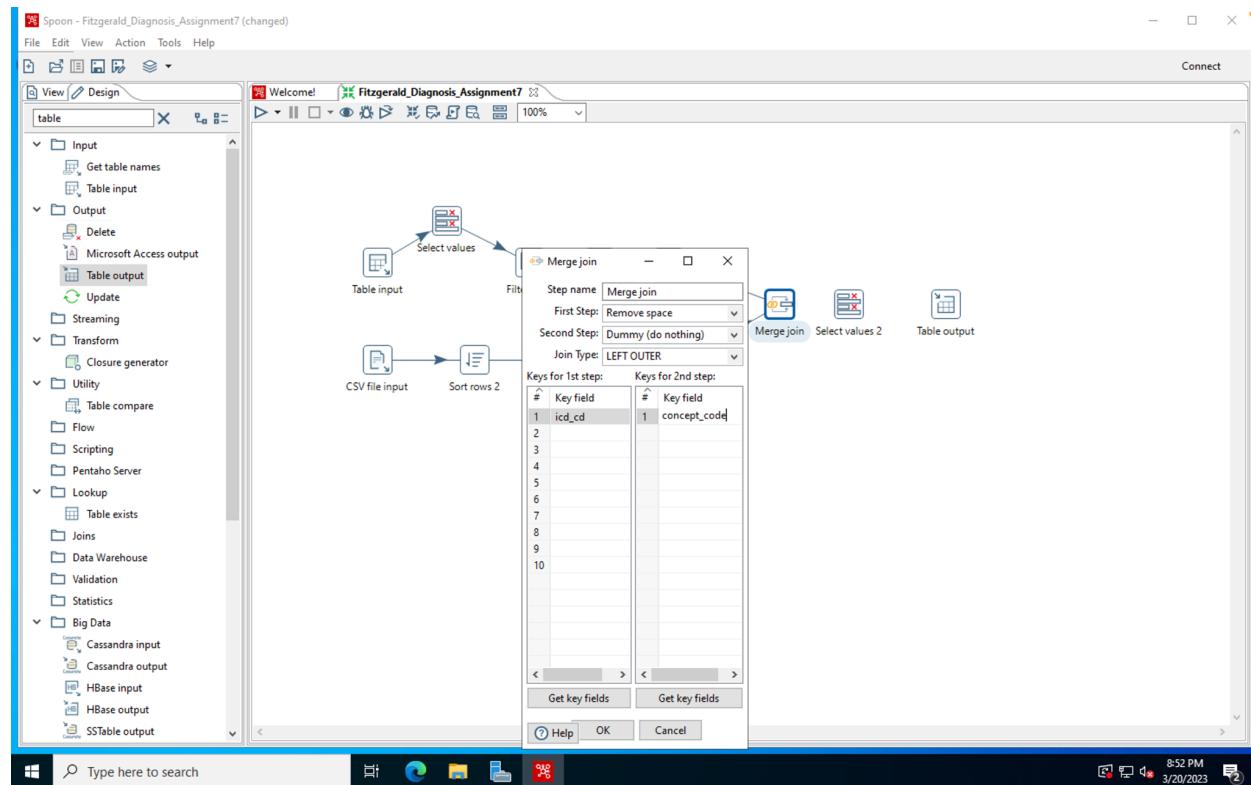


CSV Input: using Concepts.csv with ‘tab’ delimiter → Sort Rows (similar to above) → Filter Rows 2 → Dummy do nothing (when filter ‘True’)





Merge



Rename Columns

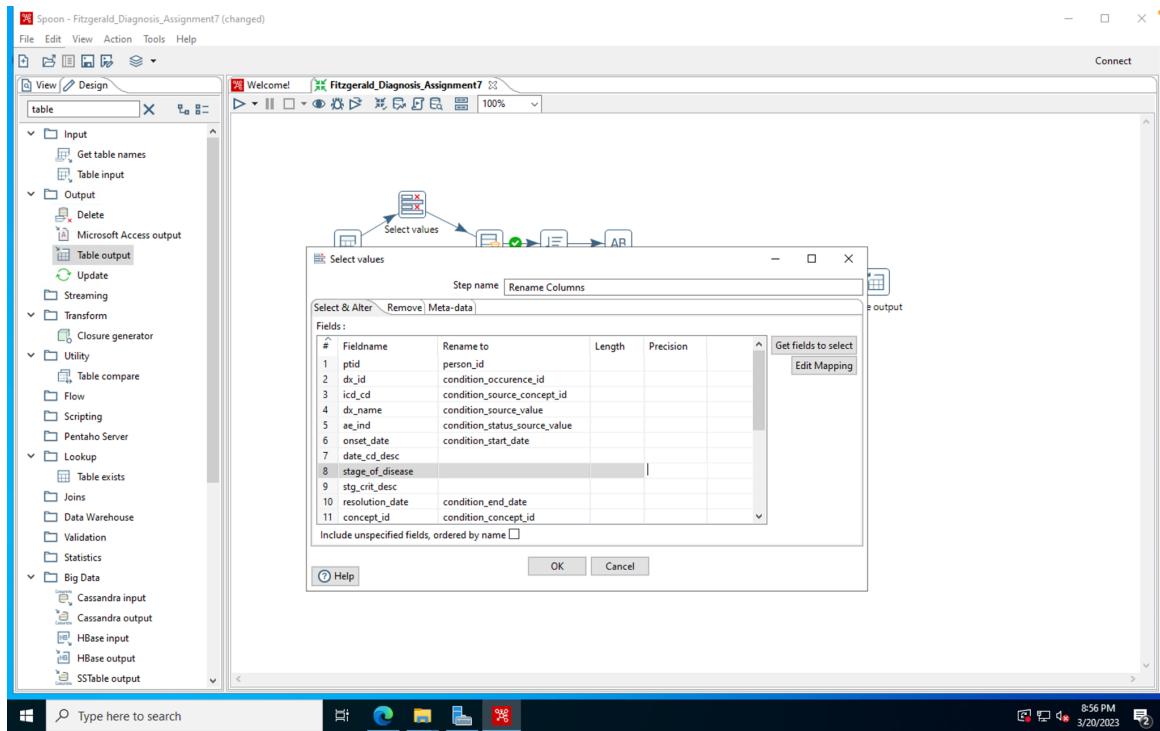
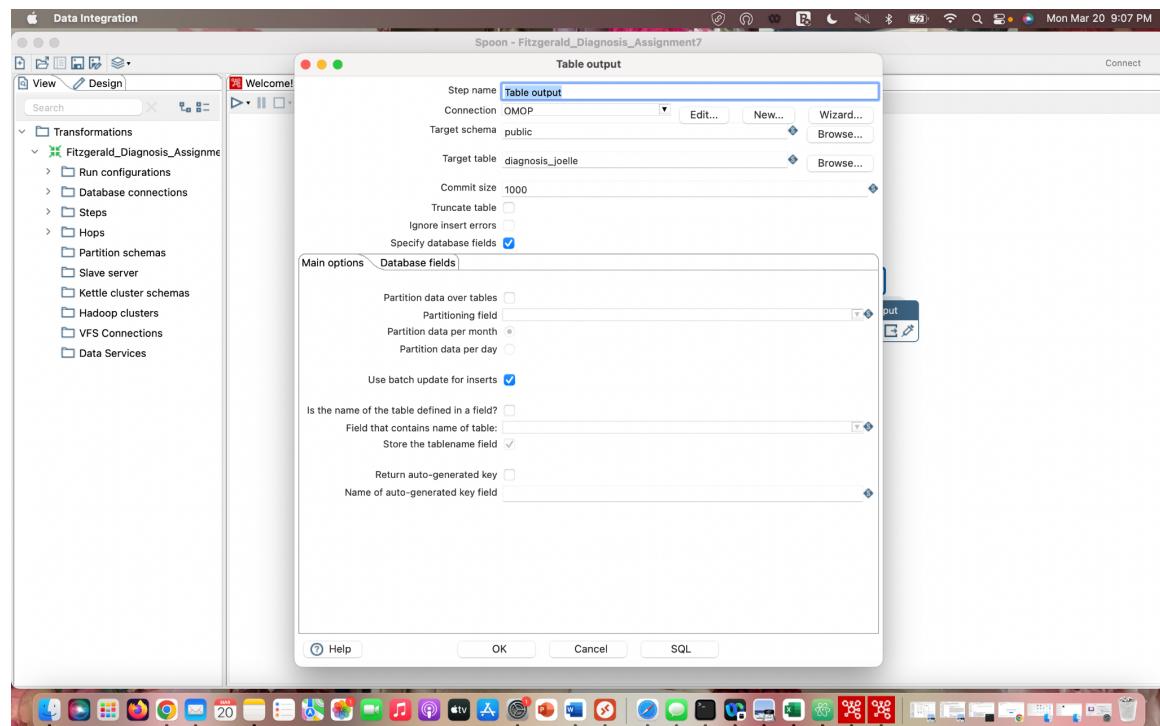


Table Output (OMOP)



Successful Mapping

